

В результате обработки получается бинарная матрица, где единицам соответствуют точки со значительным перепадом яркости, нулям – все остальные.

В дальнейшем при сегментации изображения в бинарной матрице единицами представлены точки, принадлежащие искусственно утолщенным на предыдущем этапе границам объектов. Для выделения границы одного объекта в матрице по определенному алгоритму отыскивается элемент, равный единице, не отнесенный ранее ни к какому другому объекту. На следующем шаге считается, что все соседние элементы, равные единице, также принадлежат этому объекту. Для выделения точек внешнего контура используется обход полученного объекта по внешней его стороне, начиная с нижней левой точки объекта и заканчивая ею же. В результате получаем массив точек, образующий замкнутый контур объекта. Контур [2] – граница объекта, которая представляет собой замкнутую последовательность точек (x_s, y_s) , где $1 \leq s \leq N$. Из него равномерно выбирается 128 точек (x_s, y_s) , где $1 \leq s \leq 128$, которые используются для вычисления предназначенных для индексирования характеристик формы.

Кластерный анализ. После применения описанных выше методов осуществляется кластеризация схожих изображений. Задача кластерного анализа заключается в том, чтобы на основании данных о схожести изображений разбить их множество G на m кластеров (подмножеств) Q_1, Q_2, \dots, Q_m , так, чтобы каждый объект G_j принадлежал одному и только одному подмножеству разбиения и чтобы объекты, принадлежащие одному и тому же кластеру, были сходными, в то время как объекты, принадлежащие разным кластерам, были разнородными.

Заключение. Таким образом, применяя различные методы сравнения изображений, можно найти похожие изображения в коллекции графических файлов. Группируя найденные изображения в кластеры, становится возможной работа с группами похожих изображений внутри большой группы файлов. Внутри определенного кластера, размеры которого во много раз меньше общей совокупности графических файлов коллекции, возможна дальнейшая обработка.

Список литературы: 1. Байгарова Н.С., Бухиштаб Ю.А., Евтеева Н.Н., Корягин Д.А. Некоторые подходы к организации содержательного поиска изображений и видеoinформации.– Москва: Институт прикладной математики им. Келдыша РАН, 2002.– 24 с. 2. Гончаров А., Мельниченко А. Pseudometric Approach to Content Based Image Retrieval and Near Duplicates Detection.– РОМИП, 2008.– 34 с. 3. Пономаренко Н.Н., Лукин В.В., Абрамов С.К. Устойчивый поиск изображений по полному и тематическому подобию с использованием многопараметровой классификации.– Харьков: «ХАИ», 2003.– 13с.

Поступила в редакцию 27.11.09

УДК 621.3

Н. И. БЕЗМЕНОВ, канд. техн. наук, профессор НТУ «ХПИ»,
С. В. КОВАЛЕНКО, старший преподаватель НТУ «ХПИ»,
М. А. ТКАЧЕВ, магистрант НТУ «ХПИ»

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ ГРАФИЧЕСКИХ ОБРАЗОВ

Пропонується використовувати нейронні мережі на основі базисно радіальних функцій як метод для розпізнавання графічних образів. Продемонстрована працездатність мережі на прикладі розпізнавання цифр.

Предлагается использовать нейронные сети на основе базисно радиальных функций в качестве метода для распознавания графических образов. Продемонстрирована работоспособность сети на примере распознавания цифр.

It is proposed to use neural network based on radial basic functions as a method for recognition of graphic images. The capacity of network is rotined on the example of recognition of numbers.

Введение. Одной из актуальных проблем современных информационных технологий в настоящее время остается распознавание образов, в частности, графически представленных символов. Для решения данной задачи используют ряд алгоритмов, которые можно условно разделить на три категории:

- алгоритмы, основанные на сравнении шаблонов;
- алгоритмы, основанные на методах теории решений;
- алгоритмы, использующие нейронные сети.

Все вышеперечисленные группы алгоритмов имеют, в целом, одинаковую структуру и содержат такие блоки:

- блок обучения системы на наборе образов, подобным тем, которые будут распознаваться системой в будущем. По завершению обучения, его результаты сохраняются;
- блок распознавания, который, используя информацию, полученную на этапе обучения, относит незнакомый образ к одному из классов с помощью определенного математического аппарата (классификатора).

Алгоритмы, основанные на сравнении шаблонов, из-за своей негибкости (привязанности к шрифту, масштабу, наклону символов) не подходят для общего решения данной задачи. Однако в силу своей простоты и малым вычислительным затратам, они (с определенными модификациями) достаточно часто используются при решении отдельных задач распознавания [1].

Метод нейронных сетей, использующий некоторые принципы искусственного интеллекта, широко применяется в большинстве

современных систем распознавания символьной информации. Среди его преимуществ стоит упомянуть универсальность и незначительные требования к проектированию алгоритмов. Недостатками данного метода являются значительный объем вычислительных ресурсов, необходимых для организации процесса обучения, а также плохая предсказуемость результатов распознавания. Наибольшим преимуществом таких методов является высокая точность распознавания.

Алгоритмы, основанные на методах теории решений, во много раз упрощают процесс обучения, но их точность значительно уступает точности нейросетевых алгоритмов. Важным преимуществом данных алгоритмов является представление символов в векторной форме (в виде набора характеристик), в то время как большинство нейросетевых алгоритмов оперируют с растровым изображением. При использовании некоторых классификаторов теории решений, имеется возможность добиться высокой степени предсказуемости результатов распознавания.

Далее рассматривается метод распознавания графических образов на основе нейронных сетей. В качестве архитектуры нейросети была выбрана сеть, основанная на радиально базисных функциях.

Введение в RBF-сети. Базовая архитектура сетей на основе радиальных базисных функций (radial basis function network - RBF), или RBF-сетей, предполагает наличие двух слоев, выполняющих совершенно различные функции. Входной слой состоит из сенсорных элементов, которые связывают сеть с внешней средой. Второй слой является единственным скрытым слоем сети и выполняет нелинейное преобразование входного пространства в скрытое.

Математическую основу функционирования радиальных сетей составляет теорема Т. Ковера о распознаваемости образов [3], в соответствии с которой нелинейные проекции образов в некоторое многомерное пространство могут быть линейно разделены с большей вероятностью, чем при их проецировании в пространство с меньшей размерностью.

Простейшая нейронная сеть радиального типа функционирует по принципу многомерной интерполяции, состоящей в отображении p различных входных векторов $x_i, i=1,2,\mathbf{K},p$ из входного N -мерного пространства во множество из p чисел $d_i, i=1,2,\mathbf{K},p$. Для реализации этого процесса необходимо использовать p скрытых нейронов радиального типа и задать такую функцию отображения $F(x)$, для которой выполняется условие интерполяции $F(x_i) = d_i$.

Использование p скрытых нейронов, соединяемых связями с весами с выходными линейными нейронами, означает формирование выходных сигналов сети путем суммирования взвешенных значений соответствующих базисных функций. Рассмотрим радиальную сеть с одним выходом и p

обучающими парами (x_i, d_i) и будем считать, что координаты каждого из p центров узлов сети определяются одним из векторов x_i , т.е. $c_i = x_i$. В этом случае взаимосвязь между входными и выходными сигналами сети может быть определена системой уравнений, линейных относительно весов. Эта система в матричной форме имеет вид:

$$j \cdot w = d, \quad (1)$$

где j_{ji} – определяет радиальную функцию с центром в точке x_i с вынужденным вектором x_j , $w = [w_1, w_2, \mathbf{K}, w_p]^T$ и $d = [d_1, d_2, \mathbf{K}, d_p]^T$.

Доказано [2], что для ряда радиальных функций в случае $x_1 \neq x_2 \neq \mathbf{K} x_p$ квадратная интерполяционная матрица j является невырожденной и при этом неотрицательно определенной. Поэтому существует решение уравнения (1) в виде:

$$w = (j^T j)^{-1} j^T d, \quad (2)$$

что позволяет получить вектор весов выходного нейрона сети.

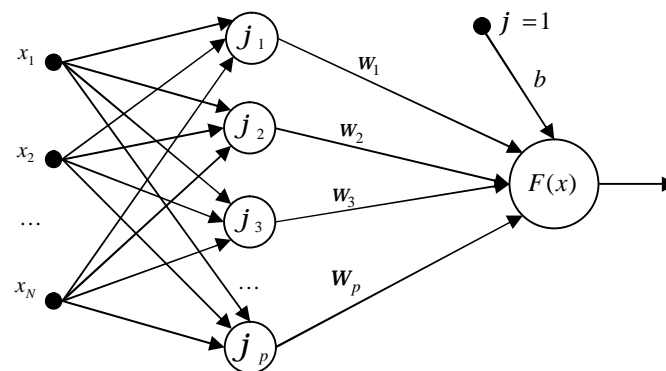


Рис. 1. Обобщенная структура радиальной сети

Чаще всего в качестве радиальной функции применяется функция Гаусса. При размещении ее центра в точке c_i она может быть определена в сокращенной форме:

$$j(x) = j(\|x - c_i\|) = \exp\left(\frac{-\|x - c_i\|^2}{2s_i^2}\right). \quad (3)$$

В этом выражении S_i - параметр, от значения которого зависит ширина функции.

На рис. 1 в общем виде изображена структура RBF-сети. Это сеть с двухслойной структурой, в которой только скрытый слой выполняет нелинейное отображение, реализуемое нейронами с базисными радиальными функциями. Выходной нейрон, как правило, линеен, а его роль сводится к взвешенному суммированию сигналов, поступающих от нейронов скрытого слоя.

Пример работы сети. В качестве примера работы сети попытаемся различить символы «5» и «4», изображения которых представлены на рис. 2.

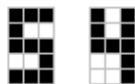


Рис. 2. Тестовые изображения

В этом случае сеть принимает вид, представленный на рис. 1. Имеется всего два образа. Количество нейронов скрытого слоя будет равняться двум. Количество сенсорных элементов, составляющих входящего вектора образа, равняется количеству элементов в изображении – 15 в данной задаче.

Запишем пару векторов обучающей выборки, воспользовавшись следующей методикой: переводя изображения в вектора (построчно перемещаясь по квадратам), будем ставить в соответствие заполненной области единицу, а пустой – 0. В результате получаем такие векторы:

$$x_1 = \{1,1,1,1,0,0,1,1,1,0,0,1,1,1\}, \quad (4)$$

$$x_2 = \{1,0,1,1,0,1,1,1,1,0,0,1,0,0,1\}. \quad (5)$$

В качестве центров примем соответственно значения $c_1 = x_1$ и $c_2 = x_2$, в результате чего получим следующий вектор d :

$$d = [0,1]^T. \quad (6)$$

Каждый его элемент является желаемым откликом сети для соответствующего вектора образа.

Искомый вектор w имеет вид $w = [w_1, w_2, b]^T$, где w_1 и w_2 – весовые коэффициенты соответствующих образов, b – порог выходного нейрона.

Пусть

$$j_{ji} = \langle \|x_j - c_i\| \rangle, j = 1,2; i = 1,2. \quad (7)$$

Тогда, подставляя векторы обучающих выборок (4) и (5) в (7), получаем матрицу j :

$$j = \begin{bmatrix} 1 & 0.018 & 1 \\ 0.018 & 1 & 1 \end{bmatrix}. \quad (8)$$

Теперь, подставляя матрицу j вида (8) и вектор желаемых откликов d вида (6) в (2), получим веса w :

$$w = \begin{bmatrix} -0.416 \\ 0.811 \\ 0.27 \end{bmatrix}, \quad (9)$$

что и завершает построение архитектуры сети.

Функция активации для данной задачи имеет следующий вид:

$$F(x) = \sum_{i=1}^2 w_i j_i (\|x - c_i\|) + b. \quad (10)$$

В качестве проверки работоспособности сети, попытаемся получить желаемые отклики на соответствующие образы. Подставляя векторы образов x_1 и x_2 в функцию активации (10), получим соответственно значения $[-0.131, 1.073]$, которые с некоторой погрешностью говорят о способности сети к классификации образов.

Выводы. Таким образом, в результате проведенных исследований, можно утверждать о способности сети к классификации векторов образов на основе базисно радиальных функций. Тестовый пример показывает положительные результаты распознавания символов печатного текста.

Список литературы: 1. Ватолин Д.С. Алгоритмы сжатия изображений. – МГУ, 2002. – С. 30- 32. 2. Micchelli C.A. "Interpolation of scattered data: Distance matrices and conditionally positive definite functions" // Constructive Approximation, 1986, vol. 2, p. 11-22. 3. Powell M.J.D. "Radial basis functions for multivariable interpolation: A review" // IMA Conference on Algorithms for the Approximation of Functions and Data, 1985, p. 143-167, RMCS, Shrivenham, England.

Поступила в редколлегию 11.12.09